

Generalized Dyson Brownian motion, McKean-Vlasov equation and eigenvalues of random matrices

Songzi Li, Xiang-Dong Li*, Yong-Xiao Xie

March 7, 2013

Abstract

Using Itô's calculus and the mass optimal transportation theory, we study the generalized Dyson Brownian motion (GDBM) and the associated McKean-Vlasov evolution equation with an external potential V . Under suitable condition on V , we prove the existence and uniqueness of strong solution to SDE for GDBM. Standard argument shows that the family of the process of empirical measures $L_N(t)$ of GDBM is tight and every accumulative point of $L_N(t)$ in the weak convergence topology is a weak solution of the associated McKean-Vlasov evolution equation, which can be realized as the gradient flow of the Voiculescu free entropy on the Wasserstein space over \mathbb{R} . Under the condition $V'' \geq -K$ for some constant $K \geq 0$, we prove that the McKean-Vlasov equation has a unique solution $\mu(t)$ and $L_N(t)$ converges weakly to $\mu(t)$ as $N \rightarrow \infty$. For C^2 convex potentials, we prove that $\mu(t)$ converges to the equilibrium measure μ_V with respect to the W_2 -Wasserstein distance on $\mathcal{P}_2(\mathbb{R})$ as $t \rightarrow \infty$. Under the uniform convexity or a modified uniform convexity condition on V , we prove that $\mu(t)$ converges to μ_V with respect to the W_2 -Wasserstein distance on $\mathcal{P}_2(\mathbb{R})$ with an exponential rate as $t \rightarrow \infty$. Finally, we discuss the double-well potentials and raise some conjectures.

Key words and phrases: Generalized Dyson Brownian motion, McKean-Vlasov equation, Johansson's theorem, gradient flow, Voiculescu free entropy, Wasserstein distance.

1 Introduction

1.1 Background

Let $V : \mathbb{R} \rightarrow \mathbb{R}^+$ be a real polynomial of even degree with positive leading coefficient, or more general a real analytic function. Consider the following probability measure on \mathcal{H}_N (the set of all $N \times N$ Hermitian matrices)

$$d\mu_N(M) = \frac{1}{Z_N} \exp(-N\text{Tr}V(M))dM,$$

*Research supported by NSFC No. 10971032, Key Laboratory RCSDS, CAS, No. 2008DP173182, and a Hundred Talents Project of AMSS, CAS.

where Z_N is a normalization constant, and if we denote x_1, \dots, x_N the eigenvalues of M ,

$$\mathrm{Tr}V(M) = \sum_{i=1}^N V(x_i).$$

By [7, 20], the distribution of the eigenvalues of M has the following probability density

$$\rho_N(x) = \frac{1}{Z_N} \prod_{i < j} |x_i - x_j|^2 \exp \left(-N \sum_{i=1}^N V(x_i) \right), \quad x \in \mathbb{R}^N.$$

The probability distribution $\rho_N(x)dx$ has a statistical mechanical interpretation: it is the canonical Gibbs measure, at inverse temperature $\beta = 2$, of a system of N unit charges interacting through the logarithmic Coulomb potential and confined by an external potential NV . From the statistical mechanical point of view, it is natural to consider the logarithmic Coulomb gas at arbitrary values of the inverse temperature $\beta > 0$. More generally, let $V : \mathbb{R} \rightarrow \mathbb{R}$ be a polynomial of even degree or an analytic function with the growth condition

$$V(x) \geq (1 + \delta) \log(x^2 + 1), \quad x \in \mathbb{R}, \quad (1)$$

the β -invariant random matrix ensemble, or the so-called log-gas model, is defined as an interacting particle system with the following probability distribution

$$P_\beta^N(dx_1, \dots, dx_N) = \frac{1}{Z_N^\beta} \prod_{i \neq j} |x_i - x_j|^{\frac{\beta}{2}} \exp \left(-\frac{\beta N}{2} \sum_{i=1}^N V(x_i) \right) \prod_{i=1}^N dx_i,$$

where $\beta > 0$ is a parameter, called the inverse of the temperature in statistical physics. When $\beta = 1, 2, 4$, one can realize the above distribution as the distribution of the eigenvalues of $N \times N$ random matrices. More precisely, for $\beta = 1, 2, 4$, if M is a $N \times N$ matrix with distribution

$$dP_\beta^N(M) = \frac{1}{Z_N} \exp \left(-\frac{\beta N}{2} \mathrm{Tr}V(M) \right) dM,$$

where dM denotes the standard measure on \mathcal{H}_N , i.e.,

$$dM = \prod_{1 \leq i \leq N} dM_{ii} \prod_{1 \leq i < j \leq N} d\mathrm{Re}(M_{ij}) d\mathrm{Im}M_{ij}.$$

In particular, for $V(x) = \frac{x^2}{2}$, we get Gaussian Symmetric matrices ensemble (GOE) for $\beta = 1$, the Gaussian Hermitian matrices ensemble (GUE) for $\beta = 2$, and for the Gaussian Symplectic matrices ensemble (GSE) for $\beta = 4$.

The distribution of interacting particles with general external potential V and the logarithmic Coulomb interaction has received a lot of attention in theoretic physics in connection with the so-called matrix models, see e.g. [9, 17]. We would like also to mention that, Kontsevich [21] also used the complex partition function for the matrix model with $V(x) = x^3$ on Hermitian random matrices to prove Witten's conjecture in the intersection theory of the moduli space of curves.

Suppose that $V : \mathbb{R} \rightarrow \mathbb{R}$ satisfies (1). In [20], Johansson proved the following result: There is a unique equilibrium measure $\mu_V \in \mathcal{P}(\mathbb{R})$ with compact support such that

$$\inf_{\mu \in \mathcal{P}(\mathbb{R})} \Sigma_V[\mu] = \Sigma_V[\mu_V],$$

and satisfies the Euler-Lagrange equation

$$\text{P.V.} \int_{\mathbb{R}} \frac{d\mu_V(y)}{x-y} = \frac{1}{2} V'(x), \quad x \in \text{supp} \mu_V,$$

where $\mathcal{P}(\mathbb{R})$ is the set of all Bore probability measures on \mathbb{R} , and Σ_V is the following energy functional, the so-called Voiculescu free entropy functional

$$\Sigma_V(\mu) = - \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x-y| d\mu(x) d\mu(y) + \int_{\mathbb{R}} V(x) d\mu(x).$$

Moreover, as N tends to infinity, the expectation of the empirical measure $L_N = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i}$ weakly converges to μ_V , i.e., $\mathbb{E}[L_N] \rightarrow \mu_V$. This recovers Wigner's famous semi-circle law [42] for GUE with $V(x) = \frac{x^2}{2}$.

Let us consider the Gaussian Hermitian Ensemble (GUE). Let B_t be the $N \times N$ Hermitian matrix with entries $B_{ij}(t)$, where $B_{ij}(t)$, $1 \leq i \leq j \leq N$, are i.i.d. complex Brownian motions with $\mathbb{E}[B_{ij}(t)] = 0$ and $\mathbb{E}[|B_{ij}(t)|^2] = t$. Let $\lambda_1(t), \dots, \lambda_N(t)$ be the process of the eigenvalues of B_t . In [15], Dyson proved that $\lambda_i(t)$ satisfies the following stochastic differential equations

$$d\lambda_N^i(t) = \frac{1}{\sqrt{N}} dW_t^i + \frac{1}{N} \sum_{j:j \neq i} \frac{1}{\lambda_N^i(t) - \lambda_N^j(t)} dt.$$

See also Mehta [28], Guionnet [18], Andersson-Guionnet-Zeitouni [2] and Tao [38].

The Dyson Brownian motion $\{\lambda_i(t), i = 1, \dots, N\}$ is an interacting N -particle system with the logarithmic Coulomb interaction. It has been very useful in various branches of mathematics and physics, including statistical physics and the quantum chaotic systems. See the references mentioned in [28, 18, 2, 34]. Instead of using Hermitian matrix-valued Brownian motion, Chan [12], Rogers and Shi [36] proved that the eigenvalues of the $N \times N$ Hermitian Ornstein-Uhlenbeck process satisfies the following stochastic differential equations

$$d\lambda_N^i(t) = \frac{1}{\sqrt{N}} dW_t^i + \frac{1}{N} \sum_{j:j \neq i} \frac{1}{\lambda_N^i(t) - \lambda_N^j(t)} dt - \frac{1}{2} \lambda_N^i(t) dt.$$

They [12, 36] proved the tightness of the family of the empirical measure $L_N(t) = \frac{1}{N} \sum_{i=1}^N \lambda_i(t)$ for the above Dyson-Ornstein-Uhlenbeck Brownian motion $(\lambda_1(t), \dots, \lambda_N(t))$ and proved that the limit of any weakly convergence subsequence of $L_N(t)$ is a weak solution of the following McKean-Vlasov equation

$$\frac{d}{dt} \int_{\mathbb{R}} f(x) \mu_t(dx) = \frac{1}{2} \int \int_{\mathbb{R}^2} \frac{\partial_x f(x) - \partial_y f(y)}{x-y} \mu_t(dx) \mu_t(dy) - \frac{1}{2} \int_{\mathbb{R}} x f'(x) \mu_t(dx),$$

where f is a test function in $C_b^2(\mathbb{R}^N)$. In particular, the Hilbert-Stiejes transformation of $\mu(t)$, defined by

$$G_t(z) = \int_{\mathbb{R}} \frac{\mu_t(dx)}{z - x},$$

satisfies the following nonlinear Burgers type equation

$$\frac{\partial}{\partial t} G_t(z) = (-G_t(z) + z) \frac{\partial}{\partial z} G_t(z) + G_t(z).$$

Showing that the above complex Burgers equation has a unique solution, they derived that $L_N(t)$ converges weakly to $\mu(t)$. Moreover, as $t \rightarrow \infty$, $G_t(z)$ converges to the Hilbert-Stiejes transform of the semi-circle law, one can therefore give a new proof of Wigner's theorem [42] for the convergence of the empirical measure of the eigenvalues of Gaussian Unitary Ensemble (GUE) to the semi-circle law. See also Guionnet [18] and Andersson-Guionnet-Zeitouni [2] for a nice presentation of this dynamic approach to Wigner's theorem using Dyson's Brownian motion.

1.2 Motivation

The purpose of this paper is to study the generalized Dyson Brownian motion and related McKean-Vlasov equation associated with the log-gas model with non-quadratic external potential. Thus, we are working on non Gaussian type β -invariant ensembles with a general external potential V . To describe our motivation, let us first introduce the generalized Dyson Brownian motions (briefly, GDBM) as follows.

Let (W^1, \dots, W^N) be an N -dimensional Brownian motion defined on a probability space (Ω, \mathbb{P}) with a filtration $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$. Let $\lambda_N(0) = (\lambda_N^1(0), \dots, \lambda_N^N(0)) \in \Delta_N$, where

$$\Delta_N = \{(x_i)_{1 \leq i \leq N} \in \mathbb{R}^N : x_1 < x_2 < \dots < x_N\}.$$

By Theorem 1.2 below, for a wide class of potential function V with suitable condition, the following stochastic differential equations

$$d\lambda_N^i(t) = \sqrt{\frac{2}{\beta N}} dW_t^i + \frac{1}{N} \sum_{j: j \neq i} \frac{1}{\lambda_N^i(t) - \lambda_N^j(t)} dt - \frac{1}{2} V'(\lambda_N^i(t)) dt, \quad i = 1, \dots, N, \quad (2)$$

have a unique strong solution $(\lambda_N(t))_{t \geq 0}$ with initial data $\lambda_N(0)$ such that $\lambda_N(t) \in \Delta_N$ for all $t \geq 0$.

The process $(\lambda_N(t))_{t \geq 0}$ defined by SDE (2) is called the generalized Dyson Brownian motion (GDBM) with potential V . The GDBM is an interacting particle system with Hamiltonian of the form

$$H(x_1, \dots, x_N) := -\frac{1}{2N} \sum_{1 \leq i \neq j \leq N} \log |x_i - x_j| + \frac{1}{2} \sum_{i=1}^N V(x_i).$$

The infinitesimal generator of these interacting particles is given by

$$\mathcal{L}_N^\beta f(x) = \frac{1}{\beta N} \sum_{k=1}^N \frac{\partial^2 f(x)}{\partial x_k^2} + \sum_{k=1}^N \left(\text{P.V.} \int_{\mathbb{R}} \frac{L_N(dy)}{x_k - y} - \frac{1}{2} V'(x_k) \right) \frac{\partial f(x)}{\partial x_k},$$

where $f \in C^2(\mathbb{R}^N)$ and $L_N = \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \in \mathcal{P}(\mathbb{R})$.

In particular, when $V \equiv 0$, $(\lambda_N(t))_{t \geq 0}$ is the Dyson Brownian motion introduced by F. Dyson [14, 15], and when $V(x) = \frac{x^2}{2}$, it is the Dyson-Ornstein-Uhlenbeck Brownian motion introduced by Chan [12] and Rogers and Shi [36]. In fact, as the same situation as what has been done by Dyson for $V = 0$, and by Chan [12] and Rogers-Shi [36] for $V(x) = \frac{x^2}{2}$, we can prove that, in the case $\beta = 1, 2, 4$, the generalized Dyson Brownian motion can be realized as the process of the eigenvalues of a matrix-valued diffusion process. See Remark 1.6, Section 2.2 and Section 2.3 below.

The process of the empirical measures of GDBM is defined by

$$L_N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_N^i(t)} \in \mathcal{P}(\mathbb{R}), \quad t \in [0, \infty).$$

In Section 3, we will prove that, under suitable condition on V , the family of the process of empirical measures $L_N(t)$ is tight, and any weak convergence limit of convergent subsequence of $L_N(t)$, denoted by $\mu(t)$, is a weak solution to the following nonlinear McKean-Vlasov equation: for all $f \in C_b^2(\mathbb{R})$,

$$\frac{d}{dt} \int_{\mathbb{R}} f(x) \mu_t(dx) = \frac{1}{2} \int \int_{\mathbb{R}^2} \frac{\partial_x f(x) - \partial_y f(y)}{x - y} \mu_t(dx) \mu_t(dy) - \frac{1}{2} \int_{\mathbb{R}} V'(x) f'(x) \mu_t(dx).$$

This also proves the existence of weak solutions of the McKean-Vlasov equation. If one can further prove the uniqueness of the weak solution to the McKean-Vlasov equation, we can then derive that the empirical measure $L_N(t)$ converges weakly to $\mu(t)$ as $N \rightarrow \infty$.

1.3 Interacting particle system

The GDBM is a special example of interacting particle systems which can be defined by a stochastic differential equation of the form

$$dx_t^i = dB_t^i - \nabla V(x_t^i) dt - \frac{1}{N} \sum_{j \neq i} \nabla W(x_t^i - x_t^j) dt, \quad i = 1, \dots, N,$$

where V is a external potential, W is an interacting function between particles, and B_t^i are i.i.d. Brownian motions on \mathbb{R}^d . See [37, 5, 11, 27, 39, 40] and reference therein. For $N = \infty$, see e.g. [22, 31]. Let

$$\mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{x_t^i}$$

be the empirical measure of the particle system $\{x_i(t), i = 1, \dots, N\}$. Then the above SDE can be rewritten as

$$dx_t^i = dB_t^i - \nabla V(x_t^i) dt - \nabla(W * \mu_t^N)(x_t^i) dt.$$

Under various assumptions which require that V and W are Lipschitz, as $N \rightarrow \infty$, it has been proved that μ_t^N converges weakly to a measure-valued process μ_t on \mathbb{R} , i.e.,

$$\mu_t^N \rightarrow \mu_t,$$

which is the law of a nonlinear diffusion process on \mathbb{R} defined by

$$d\overline{X}_t = dB_t - \nabla V(\overline{X}_t)dt - \nabla(W * \mu_t)(\overline{X}_t)dt.$$

where B_t is a Brownian motion on \mathbb{R}^d . See [27]. Suppose that $\mu_t \ll dx$, then the density function $u = \frac{d\mu_t}{dx}$ satisfies the nonlinear McKean-Vlasov equation (called also the nonlinear Fokker-Planck equation in [5] etc.)

$$\partial_t u = \operatorname{div}(u \nabla(\log u + V + W * u)). \quad (3)$$

In [5], Benedetto, Caglioti, Carrillo and Pulvirenti developed the L^1 -theory of the nonlinear McKean-Vlasov equation (3) in one-dimensional granular media. Assuming that the interaction function W and the potential V are Lipschitz and convex functions, they proved that the free energy functional

$$F(u) = \int_{\mathbb{R}} u \log u dx + \int_{\mathbb{R}} V u dx + \int \int_{\mathbb{R}^2} W(x-y) u(x) u(y) dx dy$$

has a unique minimum u_∞ , and it holds that

$$\|u_t - u_\infty\|_{L^1} \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

Moreover, under the Lipschitz condition on V and W , they proved that there exists a constant $C > 0$ such that

$$W_1(\mu_t, \nu_t) \leq e^{Ct} W_1(\mu_0, \nu_0),$$

where μ_t and ν_t are solutions to the nonlinear McKean-Vlasov equation (3) with initial dates μ_0 and ν_0 respectively. See also [37, 23] and references therein.

However, the logarithmic Coulomb interaction function appeared in the distribution of the eigenvalues of β -invariant random matrices ensemble, i.e.,

$$W(x) = -\log|x|$$

is not a Lipschitz function on \mathbb{R} . Thus, it does not satisfy the Lipschitz condition required in [27, 5], and the L^1 -theory of Benedetto et al.[5] does not apply directly to (3) associated to the GDBM, even in the case of Gaussian ensembles, i.e., $V(x) = \frac{x^2}{2}$. As far as we know, the nonlinear McKean-Vlasov equation (3) with the logarithmic Coulomb interaction $W(x) = -\log|x|$ and general external potential V has not been well studied in the literature. In particular, even though the existence of weak solution of (3) can be easily derived from the McKean-Vlasov limit of any convergent subsequence of the empirical measure of the generalized Dyson Brownian motion, the problem of the uniqueness and the longtime asymptotic behavior of the solutions of the nonlinear McKean-Vlasov equation (3) remain as an open problem in the literature.

1.4 Otto's calculus and the gradient flow of the Voiculescu entropy

In [11], Carrillo, McCann and Villani used Otto's infinite dimensional differential calculus [32] on the Wasserstein space over \mathbb{R}^n to study the convergence rate problem of the nonlinear McKean-Vlasov equation (3) in the granular media. Under some growth and smoothness

assumptions of the interaction potential W and the external potential V , they proved that, if $\nabla^2 W \geq 0$ and $\nabla^2 V \geq \lambda$, or if $\nabla^2 W \geq \lambda$ and $\nabla^2 V \geq 0$, where $\lambda > 0$ is a constant, then the solution $u(t)$ of (3) converges to the equilibrium u_∞ in the W_2 -Wasserstein distance with an exponential rate, i.e.,

$$W_2(u_t, u_\infty) = O(e^{-\lambda t}).$$

Note that, for $W(x) = -\log|x|$, we have

$$\nabla^2 W(x) = \frac{1}{|x|^2}, \quad x \neq 0.$$

Thus, the logarithmic Coulomb interaction potential has a strong convexity near its singularity point $x = 0$. This suggests us to adopt the infinite dimensional calculus on the Wasserstein space initiated by Otto [32] and developed by Carrillo, McCann and Villani [11] to study the uniqueness problem and the long time asymptotic behavior of the solution of the McKean-Vlasov evolution equation.

Let V be a C^1 function on \mathbb{R} . According to Voiculescu [41], we introduce the free entropy as follows

$$\Sigma_V(\mu) = - \int \int_{\mathbb{R}^2} \log|x-y| d\mu(x) d\mu(y) + \int_{\mathbb{R}} V(x) d\mu(x).$$

It has the following electrostatic interpretation: Suppose that electrons are distributed on the one-dimensional space \mathbb{R} in the presence of an external field with potential V . For a probability distribution μ of electrons, the electrostatic repulsion is given by

$$\int \int_{\mathbb{R}^2} \log|x-y|^{-1} d\mu(x) d\mu(y),$$

and the energy from the external field is $\int_{\mathbb{R}} V(x) d\mu(x)$. Thus, the Voiculescu free entropy is the total energy of electrons in an external field with potential V . When V satisfies the growth condition (1), it is shown in [20] that there exists a unique minimizer of the Voiculescu free entropy

$$\mu_V = \operatorname{argmin}_{\mu \in \mathcal{P}(\mathbb{R})} \Sigma_V(\mu),$$

called the equilibrium distribution of Σ_V . The relative free entropy is defined as follows

$$\Sigma_V(\mu|\mu_V) = \Sigma_V(\mu) - \Sigma_V(\mu_V).$$

Following Voiculescu [41] and Biane [3], the relative free Fisher information is defined as follows

$$I_V(\mu) = \int_{\mathbb{R}} \left(H\mu(x) - \frac{1}{2} V'(x) \right)^2 d\mu(x),$$

where

$$H\mu := \operatorname{PV} \int_{\mathbb{R}} \frac{d\mu(y)}{x-y}.$$

is the Hilbert transform of μ .

Note that, as the equilibrium measure μ_V satisfies the equation

$$H\mu_V(x) = \frac{1}{2}V'(x), \quad \forall x \in \mathbb{R},$$

we have

$$I_V(\mu_V) = 0.$$

Inspired by Otto [32], Carrillo-McCann-Villani [11], Villani [39, 40] and Biane [3], we have the following result, which has been known by Biane and Speicher [4].

Theorem 1.1 (*Biane and Speicher [4]*) *Let $V \in C^2(\mathbb{R})$. Then the McKean-Vlasov equation (5) is the gradient flow of the Voiculescu free entropy Σ_V on the Wasserstein space $\mathcal{P}_2(\mathbb{R})$ over \mathbb{R} equipped with Otto's infinite dimensional Riemannian structure.*

For the definition of the Otto's infinite dimensional Riemannian structure on the Wasserstein space $\mathcal{P}_2(\mathbb{R})$, see Section 4.

1.5 Main results

We now state the main results of this paper. Our first result gives the existence and uniqueness of the strong solution to the SDE of the generalized Dyson Brownian motion under reasonable condition on the external potential.

Theorem 1.2 *Let (W^1, \dots, W^N) be an N -dimensional Brownian motion in a probability space (Ω, \mathbb{P}) equipped with a filtration $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$. Let $V \in C^2(\mathbb{R})$ be a function satisfying the growth condition (1) and the following conditions*

(i) *For all $R > 0$, there is $K_R > 0$, such that for all $x, y \in \mathbb{R}$ with $|x|, |y| \leq R$,*

$$(x - y)(V'(x) - V'(y)) \geq -K_R|x - y|^2,$$

(ii) *For some $\gamma > 0$, such that*

$$-xV'(x) \leq \gamma(1 + |x|^2), \quad \forall x \in \mathbb{R}. \quad (4)$$

Then, for any $\beta \geq 1$, and for any given $\lambda_N(0) = (\lambda_N^1(0), \dots, \lambda_N^N(0)) \in \Delta_N$, there exists a unique strong solution $(\lambda_N(t))_{t \geq 0}$ with infinite lifetime to SDE (2), with initial value $\lambda_N(0)$ and such that $\lambda_N(t) \in \Delta_N$ for all $t \geq 0$.

The second result of this paper is the following result concerning the existence and uniqueness of the weak solution to the McKean-Vlasov equation, as well as the convergence of the empirical measure $L_N(t)$ towards the solution of the McKean-Vlasov equation.

Theorem 1.3 (i) *Suppose that V be a C^2 function satisfying the same condition as in Theorem 1.2. Suppose that*

$$\sup_{N \geq 0} \int_{\mathbb{R}} \log(x^2 + 1) dL_N(0) < \infty,$$

and

$$L_N(0) = \frac{1}{N} \sum_{k=1}^N \delta_{\lambda_N^k(0)} \rightarrow \mu \in \mathcal{P}(\mathbb{R}) \quad \text{as } N \rightarrow \infty.$$

Then, the family $\{L_N(t), t \in [0, T]\}$ is tight, and the limit of any weakly convergent subsequence of $\{L_N(t), t \in [0, T]\}$ is a weak solution of the McKean-Vlasov equation, i.e., for all $f \in C_b^2(\mathbb{R})$, $t \in [0, T]$,

$$\frac{d}{dt} \int_{\mathbb{R}} f(x) \mu_t(dx) = \frac{1}{2} \int \int_{\mathbb{R}^2} \frac{\partial_x f(x) - \partial_y f(y)}{x - y} \mu_t(dx) \mu_t(dy) - \frac{1}{2} \int_{\mathbb{R}} V'(x) f'(x) \mu_t(dx). \quad (5)$$

Equivalently, the probability density of μ_t with respect to the Lebesgue measure on \mathbb{R} satisfies the nonlinear Fokker-Planck equation

$$\frac{\partial \rho_t}{\partial t} = \frac{\partial}{\partial x} \left(\rho_t \left(\frac{1}{2} V' - H \rho_t \right) \right), \quad (6)$$

where

$$H\rho(x) = \text{P.V.} \int_{\mathbb{R}} \frac{\rho(y)}{x - y} dy$$

is the Hilbert transform of ρ .

(ii) Suppose that there exists a constant $K \in \mathbb{R}$ such that

$$V''(x) \geq K, \quad x \in \mathbb{R}.$$

Let $\mu_i(t)$ be two solutions of the McKean-Vlasov equation (5) with initial datas $\mu_i(0)$, $i = 1, 2$. Then for all $t > 0$, we have

$$W_2(\mu_1(t), \mu_2(t)) \leq e^{-Kt} W_2(\mu_1(0), \mu_2(0)).$$

In particular, the McKean-Vlasov equation (5) has a unique solution.

(iii) Let V be a C^2 function satisfying the same condition as in Theorem 1.2 and $V'' \geq K$ for some constant $K \in \mathbb{R}$. Then the empirical measure $L_N(t)$ weakly converges to the unique solution $\mu(t)$ of the McKean-Vlasov equation (5).

The following result gives us the longtime asymptotic behavior and the convergence rate of the weak solution of the McKean-Vlasov equation to the equilibrium measure of the Voiculescu free entropy for C^2 -convex potentials.

Theorem 1.4 (i) Let V be a C^2 -convex potential. Then $\mu(t)$ converges to μ_V with respect to the Wasserstein distance in $\mathcal{P}_2(\mathbb{R})$, i.e.,

$$W_2(\mu(t), \mu_V) \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

(ii) Suppose that there exists a constant $K \in \mathbb{R}$ such that

$$V''(x) \geq K, \quad \forall x \in \mathbb{R}.$$

Then for all $t > 0$, we have

$$\begin{aligned} \Sigma_V(\mu_t | \mu_V) &\leq e^{-2Kt} \Sigma_V(\mu_0 | \mu_V), \\ W_2(\mu_t, \mu_V) &\leq e^{-Kt} W_2(\mu_0, \mu_V). \end{aligned}$$

In particular, if $K > 0$, then μ_t converges to μ_V with respect to the W_2 -Wasserstein distance with the exponential rate K .

(iii) Suppose that V is a C^2 , convex and there exists a constant $r > 0$ such that

$$V''(x) \geq K > 0, \quad |x| \geq r.$$

Then there exist constants $C_1 > 0$ and $C_2 > 0$ such that $\mu(t)$ converges to μ_V with respect to the W_2 -Wasserstein distance with an exponential rate

$$W_2^2(\mu_t, \mu_V) \leq \frac{e^{-C_1 t}}{C_2} \Sigma_V(\mu_0 | \mu), \quad t > 0.$$

1.6 Remarks

Remark 1.5 (i) In [36], Rogers and Shi proved (even though did not state) the existence and uniqueness of the generalized Dyson Brownian motion under the condition that the potential V satisfies

$$-xV'(x) \leq \gamma, \quad \forall x \in \mathbb{R}. \quad (7)$$

It is clear that our condition (4) on V is weaker than Rogers-Shi's condition (7). To prove Theorem 1.2, we need to use the non-explosion criterion of the Bessel type SDE on the half line \mathbb{R}^+ (cf. [19] p. 235-237), which was kindly pointed to us by Yves Le Jan in June 2011. (ii) The conclusion of Theorem 1.6 says that, for any given $\lambda_N^1(0) < \lambda_N^2(0) < \dots < \lambda_N^{N-1}(0) < \lambda_N^N(0)$, SDE (2) admits a unique strong solution with infinite lifetime such that $\lambda_N^1(t) < \lambda_N^2(t) < \dots < \lambda_N^{N-1}(t) < \lambda_N^N(t)$. Therefore, the generalized Dyson Brownian particle system $(\lambda_N^1(t), \dots, \lambda_N^N(t))$ does not self-intersect for all time $t > 0$.

Remark 1.6 Let $\beta = 1, 2, 4$. Let V be an analytic function satisfying the assumptions of Theorem 1.2. Following an original idea due to Dyson [15] and developed by other authors [12, 36, 2, 18] for the special case $V(x) = \theta x^2$, we can prove that the generalized Dyson Brownian motion $\lambda_N^1(t) \leq \dots \leq \lambda_N^N(t)$ can be realized as the eigenvalues of a matrix-valued diffusion process X_t , which satisfies the following SDE

$$dX_t = \sqrt{\frac{2}{\beta N}} dW_t - \frac{1}{2} \nabla \text{Tr} V(X_t) dt, \quad (8)$$

where B_t is the standard Brownian motion on \mathcal{H}_N^β (see Section 2.1), and ∇ denotes the gradient operator on \mathcal{H}_N^β . Moreover, $\lambda_N(t) = (\lambda_1(t), \dots, \lambda_N(t))$ is a Δ_N -valued semi-martingale. See Theorem 2.2 and Theorem 2.3.

Remark 1.7 In [3], Biane pointed out that there exists non-convex potentials such that the equation $H\rho_i(x) = \frac{1}{2}V'(x)$ on the supports of μ_i holds for distinct measures $\mu_i = \rho_i(x)dx$. The simplest example due to Biane [3] is a two-well potential V satisfying $V(x) = \frac{1}{2}(x - a_i)^2$ in $[a_i - 2, a_i + 2]$ for points a_1, a_2 with $|a_1 - a_2| > 4$, then the semi-circular measures centered on a_1 and a_2 respectively satisfy $H\rho_i(x) = V'(x)$ on their supports respectively. By Theorem 1.2 (due to Biane and Speicher [4]), the McKean-Vlasov limit $\mu(t) = \rho_t(x)dx$ of $L_N(t)$ satisfies the free Fokker-Planck equation

$$\frac{\partial \rho_t}{\partial t} = \frac{\partial}{\partial x} \left(\rho_t \left(\frac{1}{2} V' - H\rho_t \right) \right).$$

It is natural to ask the question whether $\mu(t)$ converges to μ_V in the weak convergence topology or with respect to the W_2 -Wasserstein distance (or the W_p -Wasserstein distance for some $p \geq 1$) for non-convex potentials V . In [4], it was pointed out that this cannot be true in general. Indeed, as μ_t satisfies the gradient flow of the Voiculescu free entropy Σ_V on $\mathcal{P}(\mathbb{R})$, μ_t may converge to a local minimizer of Σ_V which is not necessary the global minimizer μ_V . See also Section 6 and in particular Conjecture 6.1.

The rest of this paper is organized as follows. In Section 2, we prove Theorem 1.2. In Section 3, we prove Theorem 1.3 (i). In Section 4, we prove Theorem 1.4. In Section 5, we prove Theorem 1.3 (ii) and (iii). Finally, we study the double-well potentials and raise some conjectures in Section 6.

Acknowledgement. During the preparation of this paper, the authors have benefited from some discussions with G. Akemann, F. Götze, K. Kuwae, Y. Le Jan, Y. Liu, T. Lyons, S.-G. Peng, M. Röckner, W. Sun and M. Venker. We would like to express our gratitude to all of them for their interests and for useful comments.

2 Generalized Dyson's Brownian motion (GDBM)

2.1 Proof of Theorem 1.2

For $T \in \mathbb{R}^+$, we denote by $\mathcal{C}([0, T], \mathbb{R}^N)$ the space of continuous process from $[0, T]$ to \mathbb{R}^N and $\mathcal{P}(\mathcal{C}([0, T], \Delta_N))$ is the set of all probability measures on $\mathcal{C}([0, T], \Delta_N)$.

(1) Fix $R > 0$. Define the truncated Dyson Brownian motion by

$$d\lambda_{N,R}^i(t) = \sqrt{\frac{2}{\beta N}} dW_t^i + \frac{1}{N} \sum_{j:j \neq i} \phi_R(\lambda_N^i(t) - \lambda_N^j(t)) dt - \frac{1}{2} V'(\lambda_N^i(t)) dt, \quad (9)$$

with $\lambda_{N,R}^i(0) = \lambda_N^i(0)$ for $1 \leq i \leq N$, where $\phi_R(x) = x^{-1}$ if $|x| \geq R^{-1}$, and $\phi_R(x) = R^2 x$ if $|x| < R^{-1}$. By Theorem 3.1.1 in [35], since ϕ_R is uniformly Lipschitz and V satisfies the conditions (i) and (ii), SDE (9) has a unique strong solution, adapted to the filtration \mathcal{F} . Let

$$\tau_R := \inf\{t : \min_{i \neq j} |\lambda_{N,R}^i(t) - \lambda_{N,R}^j(t)| < R^{-1}\}.$$

Then τ_R is monotone increasing in R and

$$\lambda_{N,R}(t) = \lambda_{N,R'}(t) \quad \text{for all } t \leq \tau_R \text{ and } R < R'. \quad (10)$$

(2) We now construct a solution to SDE (2) by taking $\lambda_N(t) = \lambda_{N,R}(t)$ on the event $[\tau_R > t] = \{|\lambda_N^i(s) - \lambda_N^j(s)| \geq R^{-1}, \forall s \leq t, \forall i \neq j\}$. To obtain a global solution $\{\lambda_N(t), t \in \mathbb{R}^+\}$ to SDE (2), we need only to prove that τ_R tends to infinity almost surely when $R \rightarrow \infty$.

To this end, let us consider the Lyapounov function

$$f(x_1, \dots, x_N) = \frac{1}{N} \sum_{i=1}^N V(x_i) - \frac{1}{N^2} \sum_{i \neq j} \log |x_i - x_j|. \quad (11)$$

Using $\log |x - y| \leq \log(|x| + 1) + \log(|y| + 1)$ and (1), we can derive that

$$f(x_1, \dots, x_N) \geq -(1 + \delta) \log 2. \quad (12)$$

and $V(x) - 2 \log(|x| + 1) \geq C$ for $C := -2(1 + \delta) \log 2$. Moreover, for all $i \neq j$, it holds

$$-\frac{1}{N^2} \log |x_i - x_j| \leq f(x_1, \dots, x_N) - C. \quad (13)$$

For any $M > 0$, set

$$T_M = \inf\{t \geq 0 : f(\lambda_N(t)) \geq M\}.$$

Then, T_M is a stopping time. Moreover, on $\{T_M \geq T\}$, we have

$$|\lambda_N^i(t) - \lambda_N^j(t)| \geq R^{-1} := e^{N^2(-M+C)}, \quad \forall t \leq T.$$

Thus, on the event $[T \leq T_M]$, $\{\lambda_{N,R}(t), t \leq T\}$ is an adapted solution to SDE (2). It remains to prove that for all $t \geq 0$, we have

$$\mathbb{P}(\exists M \in \mathbb{N} : T_M \geq t) = 1. \quad (14)$$

(3) To prove (14), we need only to prove that, for all $K > 0$, we have

$$\mathbb{P}(\exists M \in \mathbb{N} : T_M \wedge \zeta_K \geq t) = 1, \quad (15)$$

where

$$\zeta_K = \inf\{t \geq 0 : \lambda_N^i(t) \notin [-K, K], \text{ for some } i = 1, \dots, N\}.$$

By (11) and (13), to prove (15), it is equivalent to show that almost surely $\lambda_N^i(t)$ and $\lambda_N^j(t)$ never collide up to ζ_M . We shall prove this claim below.

(4) By Itô's formula, for all $f \in C^2(\mathbb{R})$, we have

$$\begin{aligned} df(\lambda_N(t)) &= \frac{1}{N^2} \sum_{i=1}^N \left(V'(\lambda_N^i(t)) - \frac{1}{N} \sum_{k \neq i} \frac{1}{\lambda_N^i(t) - \lambda_N^k(t)} \right) \left(\sum_{j \neq i} \frac{1}{\lambda_N^i(t) - \lambda_N^j(t)} \right) dt \\ &\quad + \frac{1}{2N} \sum_{i=1}^N \left(-|V'(\lambda_N^i(t))|^2 + \frac{1}{N} \sum_{j \neq i} \frac{V'(\lambda_N^i(t))}{\lambda_N^i(t) - \lambda_N^j(t)} \right) dt \\ &\quad + \frac{1}{\beta N^2} \sum_{i=1}^N \left(V''(\lambda_N^i(t)) + \frac{1}{N} \sum_{j \neq i} \frac{1}{(\lambda_N^i(t) - \lambda_N^j(t))^2} \right) dt + dM_N(t), \end{aligned}$$

where M_N is the following local martingale

$$dM_N(t) = \frac{2^{\frac{1}{2}}}{\beta^{\frac{1}{2}} N^{\frac{3}{2}}} \sum_{i=1}^N \left(V'(\lambda_N^i(t)) - \frac{1}{N} \sum_{k \neq i} \frac{1}{\lambda_N^i(t) - \lambda_N^k(t)} \right) dW_t^i.$$

By [18], we have

$$\sum_{i=1}^N \left[\sum_{k \neq i} \frac{1}{\lambda_N^i(t) - \lambda_N^k(t)} \sum_{j \neq i} \frac{1}{\lambda_N^i(t) - \lambda_N^j(t)} - \sum_{j \neq i} \frac{1}{(\lambda_N^i(t) - \lambda_N^j(t))^2} \right] = 0.$$

Thus

$$\begin{aligned}
df(\lambda_N(t)) &= dM_N(t) + \frac{1}{N^3} \left(\frac{1}{\beta} - 1 \right) \sum_{k \neq i} \frac{1}{(\lambda_N^i(t) - \lambda_N^k(t))^2} dt - \frac{1}{2N} \sum_{i=1}^N |V'(\lambda_N^i(t))|^2 dt \\
&\quad + \frac{1}{N^2} \left(\frac{1}{\beta} \sum_{i=1}^N V''(\lambda_N^i(t)) + \frac{3}{2} \sum_{j \neq i} \frac{V'(\lambda_N^i(t)) - V'(\lambda_N^j(t))}{\lambda_N^i(t) - \lambda_N^j(t)} \right) dt \\
&= dM_N(t) + A_N(t) dt,
\end{aligned}$$

where, as $\beta \geq 1$, it holds

$$A_N(t) \leq \frac{1}{N^2} \left(\frac{1}{\beta} \sum_{i=1}^N V''(\lambda_N^i(t)) + \frac{3}{2} \sum_{j \neq i} \frac{V'(\lambda_N^i(t)) - V'(\lambda_N^j(t))}{\lambda_N^i(t) - \lambda_N^j(t)} \right) dt.$$

We now prove the following lemma, which is a stronger version of the claim in **(3)**.

Lemma 2.1 *Suppose that the processes $(\lambda_N(t))_{t \geq 0}$ defined by (2), at least up to the stopping time*

$$\zeta = \inf\{t : \lambda_N^i(t) = \lambda_N^j(t) \exists i \neq j \text{ or } \lambda_N^i(t) = \infty \exists i\}.$$

Then

$$\mathbb{P}(\zeta = \infty) = 1.$$

Proof. Fix $K > 0$, $T > 0$ and $R > 0$ such that $\lambda_N^i(0) \in [-K, K]$ for all $i = 1, \dots, N$, and $|\lambda_N^i(0) - \lambda_N^j(0)| \geq R^{-1}$ for all $i \neq j$, $i, j = 1, \dots, N$. Let $C_1(K) \geq 0$ be such that $V''(x) \leq C_1(K)$ for all $x \in [-K, K]$. Then, $A_N(t) \leq C_1(K)$ and $\{f(\lambda_N(t \wedge \zeta_K)) - C_1(K)(t \wedge \zeta_K), t \in [0, T]\}$ is a super-martingale.

Let $A := \{\tau_R \leq \zeta_K, \tau_R \leq T\}$, and $C_2(K) := \inf\{V(x) : |x| \leq K\}$. Then

$$\begin{aligned}
&f(\lambda_N(0)) + TC_1(K) \geq \mathbb{E}(f(\lambda_N(T \wedge \zeta_K \wedge \tau_R))) \\
&= \mathbb{E}(f(\lambda_N(\tau_R))1_A) + \mathbb{E}(f(\lambda_N(T \wedge \zeta_K))1_{A^c}) \\
&\geq \left(\frac{1}{N^2} \log R - \frac{1}{N^2} (N^2 - N - 1) \log(2K) + C_2(K) \right) \mathbb{P}(A) \\
&\quad + \left(-\frac{1}{N^2} (N^2 - N) \log(2K) + C_2(K) \right) \mathbb{P}(A^c) \\
&= \left(\frac{1}{N^2} \log R + \frac{1}{N^2} \log(2K) \right) \mathbb{P}(A) - \frac{1}{N^2} (N^2 - N) \log(2K) + C_2(K),
\end{aligned}$$

whence

$$\mathbb{P}(\tau_R \leq \zeta_K, \tau_R \leq T) \leq \frac{N^2(f(\lambda_N(0)) + TC_1(K)) + N(N-1) \log(2K) - C_2(K)}{\log(2K) + \log R}.$$

Taking $R \rightarrow \infty$, for all K and T , we have $\mathbb{P}(\tau_\infty \leq \zeta_K \wedge T) = 0$. Letting T tend to infinity, we obtain $\mathbb{P}(\tau_\infty < \zeta_K) = 0$, which yields that $\sum_{i \neq j} \log |\lambda_N^i(t) - \lambda_N^j(t)| > -\infty$ almost surely

for all $t < \zeta_K$. Moreover, letting K tend to infinity, we get

$$\mathbb{P}(\tau_\infty < \zeta) = 0,$$

where $\zeta := \lim_{K \rightarrow \infty} \zeta_K = \inf\{t > 0 : f(\lambda_N(t)) = \infty\}$ is the explosion time of $f(\lambda_N(t))$. This means that the particles $\lambda_N^1(t), \dots, \lambda_N^N(t)$, could not collide before the explosion.

We now prove that ζ is infinity almost surely. To this end, let

$$R_t = \frac{1}{2N} \sum_{j=1}^N \lambda_N^j(t)^2 = \langle L_N(t), f \rangle,$$

where $f(x) = \frac{x^2}{2}$. Now $\sum_{1 \leq j \neq r \leq N} \frac{\lambda_N^j(t)}{\lambda_N^j(t) - \lambda_N^r(t)} = \frac{N(N-1)}{2}$. By Itô's formula, we have

$$\begin{aligned} dR_t &= \sqrt{\frac{2}{\beta N}} \frac{1}{N} \sum_{j=1}^N \lambda_N^j(t) dW_t^j + \frac{1}{\beta N} dt + \frac{1}{N^2} \sum_{j \neq r} \frac{\lambda_N^j(t)}{\lambda_N^j(t) - \lambda_N^r(t)} dt - \frac{1}{2} \langle L_N(t), xV'(x) \rangle dt \\ &= \sqrt{\frac{2}{\beta N}} \frac{1}{N} \sum_{j=1}^N \lambda_N^j(t) dW_t^j + \left(\frac{1}{\beta N} + \frac{N-1}{2N} - \frac{1}{2} \langle L_N(t), xV'(x) \rangle \right) dt. \end{aligned}$$

By introducing a new Brownian motion B , we have

$$dR_t = \sqrt{\frac{2}{\beta N}} \frac{1}{N} \sqrt{2NR_t} dB_t + \left(\frac{1}{\beta N} + \frac{N-1}{2N} - \frac{1}{2} \langle L_N(t), xV'(x) \rangle \right) dt.$$

Under the assumption (4), and using the comparison theorem of one dimensional SDEs, cf. [19], we can derive that

$$R_t \leq R'_t, \quad \forall t \geq 0, \quad \text{a.s.},$$

where R' solves

$$dR'_t = \frac{2}{N} \sqrt{\frac{R'_t}{\beta}} dB_t + \left(\frac{1}{\beta N} + \frac{N-1}{2N} + \frac{1}{2} \gamma + \gamma R'_t \right) dt,$$

with $R'_0 = 0$. Moreover, by Example 8.2, (8.12) in Ikeda and Watanabe [19] (p. 235-237), the process R' never explodes. So the process R cannot explode in finite time.

(5) By the continuity of the trajectory of $\lambda_N(t)$, it is easy to see that $\lambda_N(t) \in \Delta_N$ for all $t \geq 0$. The proof of Theorem 1.2 is completed. \square

2.2 From matrix diffusions to GDBM

The following result indicates a way to introduce generalized Dyson Brownian motion as the eigenvalues process of a matrix valued diffusion process. To simplify notation, let \mathcal{H}_N^β be the ensemble of $N \times N$ matrices: for $\beta = 1$, it denotes the $N \times N$ real symmetric ensemble, for $\beta = 2$, it denotes the $N \times N$ Hermitian complex ensemble, and for $\beta = 4$, it denotes the $N \times N$ symplectic ensemble.

Theorem 2.2 *Let $\beta = 1, 2, 4$, and $V : \mathbb{R} \rightarrow \mathbb{R}$ be a real analytic function. Let X_t be a \mathcal{H}_N^β -valued diffusion process defined by*

$$dX_t = \sqrt{\frac{2}{\beta N}} dB_t - \frac{1}{2} \nabla \text{Tr} V(X_t) dt, \quad (16)$$

where B_t is the standard Brownian motion on \mathcal{H}_N^β , ∇ denotes the gradient operator on \mathcal{H}_N^β . Let $\lambda_N(t) = (\lambda_N^1(t), \dots, \lambda_N^N(t))$ be the eigenvalues of $X_N(t)$. Then, $\lambda_N(t)$ satisfies the SDE (2) for the generalized Dyson Brownian motion with $\beta = 1, 2, 4$.

Proof. We only prove Theorem 2.2 for $\beta = 2$. Note that, for analytic function $V(x) = \sum_{k=0}^{\infty} a_k x^k$ on \mathbb{R} , we have

$$\nabla \text{Tr} V(X) = \sum_{k=1}^{\infty} k a_k X^{k-1} = V'(X).$$

Hence the SDE (16) for X_t can be written as follows

$$dX_t = \frac{1}{\sqrt{N}} dB_t - \frac{1}{2} V'(X) dt. \quad (17)$$

Let $x_1(t) \leq \dots \leq x_N(t)$ be the ordered eigenvalues of X_t . Let

$$f : X \rightarrow D = \text{diag}(x_1, \dots, x_N)$$

be the matrix transformation such that $X = UDU^*$, where $U = (u_1, \dots, u_N)$ is an unitary matrix. Equivalently, we have

$$Xu_i = x_i u_i,$$

i.e., u_i is the eigenvector of M with eigenvalue x_i . Write $f = (f_1, \dots, f_N)$, where $f_i(X) = x_i$, $i = 1, \dots, N$. By Itô's formula, we have

$$dx_i(t) = \nabla_{dX_t} f_i(X_t) + \frac{1}{2} \nabla^2 f_i(X_t)(dX_t, dX_t). \quad (18)$$

By the first order Hadamard variational formula, see [38], we have

$$\begin{aligned} \nabla_{dX_t} f_i(X) &= u_i^* dX_t u_i \\ &= \frac{1}{\sqrt{N}} u_i^* dB_t u_i - \frac{1}{2} u_i^* V'(X_t) u_i dt. \end{aligned}$$

Note that

$$u_i^* V'(X_t) u_i = V'(x_i(t)).$$

By the rotational invariance of the Brownian motion, $U^* B_t U$ is a Brownian motion on \mathbb{C}^N . Denote

$$W_t^i = u_i^* B_t u_i, \quad i = 1, \dots, N.$$

Then

$$\nabla_{dX_t} f_i(X) = \frac{1}{\sqrt{N}} dW_t^i - \frac{1}{2} V'(x_i(t)) dt. \quad (19)$$

On the other hand, by the second order Hadamard variational formula, see [38], it holds

$$\nabla^2 f_i(X_t)(dX_t, dX_t) = 2 \sum_{i \neq j} \frac{|u_j^* dX_t u_i|^2}{x_i(t) - x_j(t)}.$$

By Itô's calculus, we have

$$\begin{aligned}
|u_j^* dX_t u_i|^2 &= \left| \frac{1}{\sqrt{N}} u_j^* dB_t u_i - \frac{1}{2} u_j^* V'(X_t) u_i \right|^2 \\
&= \left| \frac{1}{\sqrt{N}} u_j^* dB_t u_i - \frac{1}{2} V'(x_i(t)) u_j^* u_i dt \right|^2 \\
&= \left| \frac{1}{\sqrt{N}} u_j^* dB_t u_i \right|^2 \\
&= \frac{1}{N} dt,
\end{aligned}$$

where we have used the fact that $U^* B_t U$ is a Brownian motion on \mathbb{C}^N . Hence

$$\nabla^2 f_i(X_t)(dX_t, dX_t) = \frac{2}{N} \sum_{i \neq j} \frac{1}{x_i(t) - x_j(t)} dt. \quad (20)$$

From (18), (19) and (20), we derive that $(x_1(t), \dots, x_N(t))$ satisfies the following SDE

$$dx_i(t) = \frac{1}{\sqrt{N}} dW_t^i - \frac{1}{2} V'(x_i(t)) dt + \frac{1}{N} \sum_{i \neq j} \frac{1}{x_i(t) - x_j(t)} dt. \quad (21)$$

The proof of Theorem 2.2 is completed. \square

2.3 From GDBM to matrix diffusions

The following result provides a random matrix representation for the generalized Dyson Brownian motion which is defined by solving SDE (2).

Theorem 2.3 *Let $\beta = 1, 2$ and $\lambda_N(0) \in \Delta_N$. Then, there exists a S_N (respectively, \mathcal{H}_N)-valued diffusion process $(X^{N,\beta}(t), t \geq 0)$, such that its eigenvalues process $(\lambda_N(t))_{t \geq 0}$ is a solution of the SDE (2) for the generalized Dyson Brownian motion.*

Proof. We only prove Theorem 2.3 for $\beta = 1$. Without loss of generality, we assume that $X^N(0)$ is the diagonal matrix $D = \text{diag}(\lambda_N^1, \dots, \lambda_N^N)$. Let $M > 0$ be fixed. We consider the strong solution $\lambda_N(t)$ of SDE (2) until the stopping time T_M . We let $w_{ij}, 1 \leq i < j \leq N$ be independent Brownian motions. Hereafter, all solutions will be equipped with the natural filtration $\mathcal{F}_t = \sigma(w_{ij}(s), W_i(s), s \leq t \wedge T_M)$, where W_i the Brownian motions of SDE (2), independent of $w_{ij}, 1 \leq i < j \leq N$. For $i < j$, define $R_{ij}^N(t)$ by solving SDE

$$dR_{ij}^N(t) = \sqrt{\frac{2}{N}} \frac{1}{\lambda_N^i(t) - \lambda_N^j(t)} dw_{ij}(t), \quad R_{ij}^N(0) = 0. \quad (22)$$

For $i > j$, set $R_{ij}^N(t) = -R_{ji}^N(t)$. Let O^N be the strong solution of SDE¹

$$\begin{cases} dO^N(t) = O^N(t) dR^N(t) - \frac{1}{2} O^N(t) d\langle (R^N)^T, R^N \rangle_t, \\ O^N(0) = I \end{cases}, \quad (23)$$

¹Here, for two semi-martingales A and B with values in S_N , $\langle A, B \rangle_t = (\sum_{k=1}^N \langle A_{ik}, B_{kj} \rangle_t)_{1 \leq i, j \leq N}$ is the martingale bracket of A and B , and $\langle A \rangle_t$ is the finite variation part of A at time t .

Since SDE (23) has uniformly Lipschitz coefficients, we obtain the existence and uniqueness of strong solutions of (23) with respect to the filtration \mathcal{F}_t up to the stopping time T_M . By Lemma 4.3.4 in [2], we have $O^N(t)^T O^N(t) = I$ at all times.

Let $Y^N(t) = O^N(t)^T D(\lambda_N(t)) O^N(t)$ and define $W^N(t)$ by

$$dW^N(t) = (O^N(t))^T dY^N(t) O^N(t)$$

with the initial values $W_N(0) = Y_N(0) = X_N(0)$.

By the same argument as used in [2, 18], we can prove that

$$dW_N^{ii}(t) = \sqrt{\frac{2}{N}} dW_t^i - \frac{1}{2} V'(\lambda_N^i(t)) dt, \quad \forall i = 1, \dots, N, \quad (24)$$

and for $i \neq j$,

$$dW_N^{ij}(t) = \sqrt{\frac{2}{N}} dw_{ij}(t). \quad (25)$$

Since $(O^N(t), t \geq 0)$ is adapted, $dY^N(t) = O^N(t) dW^N(t) (O^N(t))^T$ is a continuous matrix-valued semi-martingale. Set

$$\widetilde{W}_t = \begin{cases} w_{ij}(t), & \text{if } i \neq j, \\ W^i(t), & \text{if } i = j. \end{cases}$$

Then we can write $dY_N(t) = O^N(t) dW^N(t) (O^N(t))^T$ as follows

$$\begin{aligned} dY_N(t) &= O^N(t) \left(\sqrt{\frac{2}{N}} d\widetilde{W}_t - \frac{1}{2} D(V'(\lambda_N(t))) dt \right) (O^N(t))^T \\ &= \sqrt{\frac{2}{N}} O^N(t) d\widetilde{W}_t (O^N(t))^T - \frac{1}{2} O^N(t) D(V'(\lambda_N(t))) (O^N(t))^T dt. \end{aligned}$$

Set

$$B_t = \int_0^t O^N(s) d\widetilde{W}_s (O^N(s))^T.$$

By Lévy's characterization of Brownian motion, B_t is a S_N -valued Brownian motion with respect to $\mathcal{F}_t = \sigma(w_{ij}(s), W_i(s), s \leq t)$. Moreover, by the fact V is a real analytical function and $Y_N(t) = O^N(t) D(\lambda_N(t)) (O^N(t))^T$, we have

$$\nabla \text{Tr} V(Y_t) = O^N(t) D(V'(\lambda_N(t))) (O^N(t))^T.$$

Hence

$$dY_N(t) = \sqrt{\frac{2}{N}} dB_t - \frac{1}{2} \nabla \text{Tr} V(Y_t) dt. \quad (26)$$

Thus, $Y_N(t)$ is the S_N -valued diffusion process with generator $L = \frac{1}{2} \Delta_{S_N} - N \nabla \text{Tr} V \cdot \nabla$. Note that the SDE (26) is as the same type as SDE (16) that $X_N(t)$ satisfies, and $Y_N(0) = X_N(0)$. By the uniqueness of weak solution to SDE, $Y_N(t)$ has the same law as $X_N(t)$, and $\lambda_N(t)$ is the eigenvalue process of $Y_N(t)$. The proof of Theorem 2.3 is completed. \square

Remark 2.4 *Theorem 2.3 can be considered as a reverse of Theorem 2.2. The way from $X_N(t)$ to $\lambda_N(t)$ in Theorem 2.2 is more direct and does not need to deal with the question of the existence and uniqueness of strong solution of the SDE (2) with singular drift (i.e., Theorem 1.2). On the other hand, the reverse way from $\lambda_N(t)$ to $Y_N(t)$ in Theorem 2.3 shows that the generalized Dyson Brownian motion $\lambda_N(t)$ must be obtained by the way from $X_N(t)$ to $\lambda_N(t)$ as we did in Theorem 2.2. Moreover, Theorem 2.3 also shows that the SDE (2) for generalized Dyson Brownian motion has the uniqueness in distribution. We refer the reader to [30, 18, 2, 8] and references therein for further work on matrix-valued diffusion processes.*

2.4 Itô's calculus

Let (W^1, \dots, W^N) be independent Brownian motions and $(\lambda_N^1(0), \dots, \lambda_N^N(0))$ be real numbers, let $\beta \geq 1$ and let $\lambda_N(t)_{t \geq 0}$ be the unique strong solution to SDE (2). Then by Itô's calculus, we know that for all $f \in C^2([0, T] \times \mathbb{R}, \mathbb{R})$,

$$\begin{aligned} \int_{\mathbb{R}} f(t, x) L_N(t, dx) &= \int_{\mathbb{R}} f(0, x) L_N(0, dx) + \int_0^t \int_{\mathbb{R}} \partial_s f(s, x) L_N(s, dx) ds \\ &\quad + \frac{1}{2} \int_0^t \int \int_{\mathbb{R}^2} \frac{\partial_x f(s, x) - \partial_y f(s, y)}{x - y} L_N(s, dx) L_N(s, dy) ds \\ &\quad + \left(\frac{2}{\beta} - 1 \right) \frac{1}{2N} \int_0^t \int_{\mathbb{R}} \partial_x^2 f(s, x) L_N(s, dx) ds \\ &\quad - \frac{1}{2} \int_0^t \int_{\mathbb{R}} V'(x) \partial_x f(s, x) L_N(s, dx) ds + M_N^f(t), \end{aligned}$$

where M_N^f is the martingale given by

$$M_N^f(t) = \frac{1}{N} \sqrt{\frac{2}{\beta N}} \sum_{i=1}^N \int_0^t \partial_x f(s, \lambda_N^i(s)) dW_s^i, \quad \forall t \leq T.$$

Note that

$$\langle M_N^f \rangle_t = \frac{2}{\beta N^2} \int_0^t \int_{\mathbb{R}} (\partial_x f(s, x))^2 L_N(s, dx) ds \leq \frac{2 \|\partial_x f\|_{\infty}^2 t}{\beta N^2}.$$

3 McKean-Vlasov limit as $N \rightarrow \infty$

3.1 Proof of Theorem 1.3 (i)

We first prove the tightness of $\{L_N(t), t \in [0, T]\}$ in $\mathcal{C}([0, T], \mathcal{P}(\mathbb{R}))$, then we show that the limit of any convergent subsequence of $\{L_N(t), t \in [0, T]\}$ satisfies the McKean-Vlasov equation (5). Here, for $T \in \mathbb{R}^+$, we denote by $\mathcal{C}([0, T], \mathcal{P}(\mathbb{R}))$ the space of continuous processes from $[0, T]$ into $\mathcal{P}(\mathbb{R})$ equipped with the weak convergence topology.

3.1.1 The tightness of the empirical measures

We follow the argument used in [36] to prove the tightness of $\{L_N(t), t \in [0, T]\}$. Let us pick functions $f_j \in C_b^\infty(\mathbb{R}, \mathbb{C}), j = 1, 2, \dots$, which is dense in $C_b(\mathbb{R})$. Thus

$$\langle \mu, f_j \rangle = \langle \mu', f_j \rangle, \quad \forall j \Rightarrow \mu = \mu'.$$

We also pick a C^∞ function $f_0 : \mathbb{R} \rightarrow [1, \infty)$ with the properties

$$f_0(x) = f_0(-x), \quad f_0(x) \rightarrow \infty \quad \text{as } x \rightarrow \infty, \quad x \in \mathbb{R}^+.$$

Taking test functions in the Schwartz class of smooth functions whose derivatives (up to second order) are rapidly decreasing, we may assume that

$$f_j, f_j'', V'f_j' \quad \text{are uniformly bounded for all } j \geq 1.$$

By Ethier and Kurtz [16] (p.107), to prove the tightness of $\{L_N(t), t \in [0, T], N \geq 1\}$, it is sufficient to prove that for each j the sequence of continuous real-valued functions

$$\{\langle L_N(t), f_j \rangle, t \in [0, T], N \geq 1\}$$

is tight. To this end, note that, by Theorem 1.2, there is non-collision and non-explosion for the particles $\lambda_N^i(t)$ for all $t \in [0, \infty)$. By Itô's formula in Section 2.2, we have

$$\begin{aligned} d\langle L_N(t), f \rangle &= \frac{1}{N} \sqrt{\frac{2}{\beta N}} \sum_{i=1}^N f'(\lambda_N^i(t)) dW_t^i + \left\langle L_N(t), \left(\frac{2}{\beta} - 1 \right) \frac{1}{2N} f'' - \frac{1}{2} V' f' \right\rangle dt \\ &\quad + \frac{1}{2} \int \int_{\mathbb{R}^2} \frac{f'(x) - f'(y)}{x - y} L_N(t, dx) L_N(t, dy) dt. \end{aligned}$$

This yields

$$\begin{aligned} \langle L_N(t), f_j \rangle &= \langle L_N(0), f_j \rangle + \frac{1}{2} \int_0^t \int \int_{\mathbb{R}^2} \frac{f_j'(x) - f_j'(y)}{x - y} L_N(s, dx) L_N(s, dy) ds \\ &\quad - \frac{1}{2} \int_0^t \langle L_N(s), V' f_j' \rangle ds + \int_0^t \langle L_N(s), \left(\frac{2}{\beta} - 1 \right) \frac{1}{2N} f_j'' \rangle ds + M_N^{f_j}(t) \\ &= I_1(N) + I_2(N) + I_3(N) + I_4(N) + M_N^{f_j}(t), \end{aligned} \tag{27}$$

where

$$M_N^{f_j}(t) = \frac{1}{N} \sqrt{\frac{2}{\beta N}} \int_0^t \sum_{i=1}^N f_j'(\lambda_N^i(s)) dW_s^i.$$

Note that, as $L_N(0)$ is weakly convergent, $I_1(N)$ is convergent. By the assumption that f_j and f_j'' are uniformly bounded (hence f_j' are uniformly bounded), we can easily show that $\{M_N^{f_j}(t), t \in [0, T]\}$ and $I_4(N)$ converge to zero. Moreover, by the assumption that $V'f_j'$ and f_j'' are uniformly bounded, the Arzela-Ascoli theorem implies that $I_2(N)$ and $I_3(N)$ are tight in $C([0, T], \mathbb{R})$. Thus the sequence $\{(L_N(t))_{t \geq 0} : N \geq 1\}$ is tight in $C([0, T], \mathbb{R})$. Tightness also follows for $j = 0$ if we have

$$\langle L_N(0), f_0 \rangle \rightarrow \text{finite limit as } N \rightarrow \infty.$$

So let us suppose that the initial distribution $L_N(0)$ have the property $\langle L_N(0), f_0 \rangle \leq K$ for some K , for all N . For given μ_0 , we could always find $L_N(0)$ and f_0 to satisfy this and the other conditions, and this gives the tightness for $j = 0$ also.

3.1.2 Identifying of the limit process

Without loss of generality, assuming that $\{L_{N_j}(t), t \in [0, T]\}$ is a convergent subsequence in $C([0, T], \mathcal{P}(\mathbb{R}))$. Then, for all $f \in C_b^2(\mathbb{R})$, the Itô's formula (27) in Section 2.2 and the argument used in Section 3.1 show that $\langle \mu(t), f \rangle = \lim_{j \rightarrow \infty} \langle L_{N_j}, f \rangle$ satisfies the following equation

$$\begin{aligned} \int_{\mathbb{R}} f(x) \mu_t(dx) &= \int_{\mathbb{R}} f(x) \mu_0(dx) + \frac{1}{2} \int_0^t \int \int_{\mathbb{R}^2} \frac{\partial_x f(x) - \partial_y f(y)}{x - y} \mu_s(dx) \mu_s(dy) ds \\ &\quad - \frac{1}{2} \int_0^t \int_{\mathbb{R}} V'(x) f'(x) \mu_s(dx) ds. \end{aligned}$$

That is to say, $\mu(t)$ is a weak solution to the McKean-Vlasov equation (5). Suppose that μ_t is absolutely continuous with respect to the Lebesgue measure dx , and denote

$$\rho_t(x) = \frac{d\mu_t}{dx}.$$

Integrating by parts in (5), we can prove that ρ_t satisfies the nonlinear Fokker-Planck equation (6). The proof of Theorem 1.3 (i) is completed. \square

Remark 3.1 *To characterize the McKean-Vlasov limit μ_t of the family of empirical measures $\{L_N(t), t \in [0, \infty)\}$, we need only to use the test function $f(x) = (z - x)^{-1}$, where $z \in \mathbb{C} \setminus \mathbb{R}$, instead of using all test functions $f \in C_b^2(\mathbb{R})$ in the McKean-Vlasov equation (5). Let*

$$G_t(z) = \int_{\mathbb{R}} \frac{\mu_t(dx)}{z - x}$$

be the Cauchy-Stieltjes-Hilbert transform of μ_t . Then $G_t(z)$ satisfies the following equation

$$\frac{\partial}{\partial t} G_t(z) = -G_t(z) \frac{\partial}{\partial z} G_t(z) - \frac{1}{2} \int_{\mathbb{R}} \frac{V'(x)}{(z - x)^2} \mu_t(dx).$$

In particular, in the case $V(x) = \theta x^2$, since

$$- \int_{\mathbb{R}} \frac{x}{(z - x)^2} \mu_t(dx) = z \frac{\partial}{\partial z} G_t(z) + G_t(z),$$

we obtain

$$\frac{\partial}{\partial t} G_t(z) = (-G_t(z) + \theta z) \frac{\partial}{\partial z} G_t(z) + \theta G_t(z).$$

Remark 3.2 *Following [18, 2], we can also prove a stronger version of the tightness of the empirical measure. More precisely, let $T \in \mathbb{R}^+$, and assuming V satisfy the conditions in Theorem 1.6. Assume that*

$$\sup_{N \in \mathbb{N}} \int_{\mathbb{R}} \log(x^2 + 1) L_N(0, dx) < +\infty.$$

Then, for all $T \in \mathbb{R}^+$ and $L \in \mathbb{N}$, there exists a compact set $K(L) \subseteq C([0, T], \mathcal{P}(\mathbb{R}))$ such that

$$\mathbb{P}(L_N(\cdot) \in K(L)^c) \leq e^{-N^2 L}.$$

In particular, the law of $(L_N(s), s \in [0, T])$ is almost surely tight in $C([0, T], \mathcal{P}(\mathbb{R}))$. This will be used in our forthcoming paper for a study of the large deviation principle for the generalized Dyson Brownian motion. For a proof, see our forthcoming paper.

4 McKean-Vlasov equation and optimal transport

In Section 3.1, we have proved Theorem 1.3 (i), which asserts the existence of weak solution to the McKean-Vlasov equation (5), equivalently, the existence of weak solution of the nonlinear Fokker-Planck equation (6). In this section, we use the mass optimal transportation theory to study the uniqueness and the longtime asymptotic behavior of the McKean-Vlasov equation (5) and the nonlinear Fokker-Planck equation (6).

Let

$$W(x) = -2 \log |x|, \quad x \neq 0.$$

Then the nonlinear Fokker-Planck equation (6) can be rewritten as follows

$$\partial_t \rho = \nabla \cdot (\rho \nabla (V + W * \rho)). \quad (28)$$

Before going to study the nonlinear Fokker-Planck equation (6) (i.e., (28)), we first recall some results due to Carrillo, McCann and Villani. In [11], Carrillo, McCann and Villani studied the McKean-Vlasov evolution equation of the granular media, which is given by the following

$$\partial_t \rho = \nabla \cdot (\rho \nabla (\log \rho + V + W * \rho)). \quad (29)$$

They proved that the McKean-Vlasov equation can be realized as a gradient flow of a free energy functional on the infinite Wasserstein space. More precisely, we have

Theorem 4.1 (*Carillo-McCann-Villani[11]*) *Let*

$$F(f) = \int_{\mathbb{R}^d} \rho \log \rho dv + \int_{\mathbb{R}^d} \rho V dv + \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x-y) \rho(x) \rho(y) dx dy. \quad (30)$$

Then the McKean-Vlasov equation (29) is the gradient flow of F with respect to the following infinite dimensional Riemannian metric on $\mathcal{P}_2(\mathbb{R}^d)$ (cf. Otto [32]):

$$g_{fdv}(s_1, s_2) = \int_{\mathbb{R}^d} s_1 s_2 f dv,$$

where $f dv \in \mathcal{P}_2(\mathbb{R}^d)$, $s_1, s_2 \in T_{fdv} \mathcal{P}_2(\mathbb{R}^d) = \{s : \mathbb{R}^d \rightarrow \mathbb{R} : \int_M s dv = 0\}$, and

$$s_i = -\nabla \cdot (f \nabla p_i)$$

for some $p_i \in W^{1,2}(\mathbb{R}^d)$, $i = 1, 2$.

Moreover, based on Otto's infinite dimensional geometric calculation on the Wasserstein space, Carrillo, McCann and Villani [11] proved the following entropy dissipation formula

Theorem 4.2 (*Carrillo-McCann-Villani[11]*) *Denote $\xi := \nabla(\log \rho + V + W * \rho)$. Then*

$$\frac{d}{dt}F(\rho_t) = - \int_{\mathbb{R}^n} \rho |\xi|^2 dv, \quad (31)$$

$$\begin{aligned} \frac{d^2}{dt^2}F(\rho_t) &= 2 \int_{\mathbb{R}^n} \rho \text{Tr}(D\xi)^T(D\xi) dx + 2 \int_{\mathbb{R}^n} \langle D^2V \cdot \xi, \xi \rangle \rho dx \\ &\quad + \int_{\mathbb{R}^{2n}} \langle D^2W(x-y) \cdot [\xi(x) - \xi(y)], [\xi(x) - \xi(y)] \rangle d\rho(x)d\rho(y). \end{aligned} \quad (32)$$

Now let $V : \mathbb{R} \rightarrow [0, \infty)$ be a C^2 function with growth condition (1). In [4], Biane and Speicher proved that the free Fokker-Planck equation

$$\frac{\partial \rho_t}{\partial t} = - \frac{\partial}{\partial x}(\rho_t(H\rho_t - \frac{1}{2}V'))$$

is the gradient flow of Σ_V on the Wasserstein space $\mathcal{P}_2(\mathbb{R})$. See also [8].

By analogue of the proof of Theorem 4.2 in [11], and observing that for $W(x) = -2 \log |x|$ it holds

$$\xi = -\nabla(V + W * \rho) = V' - 2H\rho,$$

we can derive the following dissipation formula for the Voiculescu free entropy.

Theorem 4.3 *Let $\xi = V' - 2H\rho$. We have*

$$\frac{d}{dt}\Sigma_V(\mu_t|\mu_V) = -2 \int_{\mathbb{R}} [V'(x) - 2(H\rho)(x)]^2 \rho(x) dx, \quad (33)$$

$$\begin{aligned} \frac{d^2}{dt^2}\Sigma_V(\mu_t|\mu_V) &= 2 \int_{\mathbb{R}} V''(x) |V'(x) - 2H\rho(x)|^2 \rho(x) dx \\ &\quad + \int \int_{\mathbb{R}^2} \frac{[V'(x) - V'(y) - 2(H\rho(x) - H\rho(y))]^2}{(x-y)^2} \rho(x)\rho(y) dx dy. \end{aligned} \quad (34)$$

Proof of Theorem 1.4 (i). By Corollary 3.2 in Biane [3], for any convex V , there exists a unique equilibrium measure μ (indeed $\mu = \mu_V$) with a density ρ satisfying the Euler-Lagrange equation $H\rho(x) = \frac{1}{2}V'(x)$ for all $x \in \text{supp}(\mu)$. Thus, Σ_V has a unique minimizer μ_V . By the fact that Σ_V is lower semi continuous with respect to the weak convergence topology, see e.g. [2, 18], we see that it is also lower semi continuous with respect to the Wasserstein topology on $\mathcal{P}(\mathbb{R})$. Moreover, as V is C^2 -convex, Theorem 4.3 implies that Σ_V is a (displacement) convex functional on $\mathcal{P}_2(\mathbb{R})$.

By Proposition 4.1 in Kloeckner [24], we know that $\mathcal{P}_2(\mathbb{R})$ has vanishing sectional curvature in the sense of Alexandrov. More precisely, for any $\mu_1, \mu_2, \mu_3 \in \mathcal{P}_2(\mathbb{R})$ and for any Wasserstein geodesic $\gamma : [0, 1] \rightarrow \mathcal{P}_2(\mathbb{R})$ such that $\gamma(0) = \mu_1$ and $\gamma(1) = \mu_2$, for all $t \in [0, 1]$, it holds that

$$W_2^2(\mu_3, \gamma(t)) = tW_2^2(\mu_3, \mu_1) + (1-t)W_2^2(\mu_3, \mu_2) - t(1-t)W_2^2(\mu_1, \mu_2).$$

Therefore, $\mathcal{P}_2(\mathbb{R})$ is a nonpositively curved (NPC) space in the sense of Alexandrov (even though $\mathcal{P}_2(\mathbb{R}^n)$ is an Alexander space with nonnegative curvature for $n \geq 2$, see e.g. [1]). By Mayer [29], we can conclude that $W_2(\mu_t, \mu_V) \rightarrow 0$ holds if we only assume that V is a C^2 -convex potential. The proof of Theorem 1.4 (i) is completed. \square

Proof of Theorem 1.4 (ii). The proof follows the same argument as used in [32, 33, 11]. We use the gradient flow of the Voiculescu entropy Σ_V on the Wasserstein space and the K -convexity along the geodesic displacement between two probability measures.

Since $V'' \geq K$, we have

$$\frac{d^2}{dt^2} \Sigma_V(\mu_t | \mu_V) \geq K.$$

By Otto's calculus, we know that

$$\text{Hess} \Sigma_V(\mu_t | \mu_V) \left(\frac{\partial \mu_t}{\partial t}, \frac{\partial \mu_t}{\partial t} \right) = \frac{d^2}{dt^2} \Sigma_V(\mu_t | \mu_V),$$

which implies that

$$\text{Hess} \Sigma_V(\mu) \geq K.$$

Let $\mu(0) = \rho(0)dx$ and $\mu(1) = \rho(1)dx$ be two probability measures with compact support on \mathbb{R} , let $\mu(s) = \rho(s)dx$ be the unique geodesic in the Wasserstein space $\mathcal{P}_2(\mathbb{R})$ linking $\mu(0)$ and $\mu(1)$. Then

$$\frac{d^2}{ds^2} \Sigma_V(\mu(s)) = \text{Hess} \Sigma_V(\rho(s)) \left(\frac{\partial \rho(s)}{\partial s}, \frac{\partial \rho(s)}{\partial s} \right) \geq K \left\| \frac{\partial \rho(s)}{\partial s} \right\|_{\mathcal{P}_2(\mathbb{R})}^2.$$

Therefore, for some $\sigma^* \in (0, 1)$,

$$\begin{aligned} \Sigma_V(\rho(1)) - \Sigma_V(\rho(0)) &= \Sigma'_V(\rho(0)) + \frac{1}{2} \Sigma''_V(\rho(\sigma))|_{\sigma=\sigma^*} \\ &\geq \left\langle \frac{d\rho(\sigma)}{d\sigma}, \nabla \Sigma_V \right\rangle \Big|_{\sigma=0} + \frac{K}{2} \int_0^1 \left\| \frac{\partial \rho(\sigma)}{\partial \sigma} \right\|_{\mathcal{P}_2(\mathbb{R})}^2 d\sigma \\ &= \left\langle \frac{d\rho(\sigma)}{d\sigma}, \nabla \Sigma_V \right\rangle \Big|_{\sigma=0} + \frac{K}{2} W_2^2(\rho(0), \rho(1)). \end{aligned}$$

Similarly,

$$\Sigma_V(\rho(0)) - \Sigma_V(\rho(1)) \geq - \left\langle \frac{d\rho(\sigma)}{d\sigma}, \nabla \Sigma_V \right\rangle \Big|_{\sigma=1} + \frac{K}{2} W_2^2(\rho(0), \rho(1)).$$

Summing the two inequalities together, we obtain

$$\left\langle \frac{d\rho(\sigma)}{d\sigma}, \nabla \Sigma_V \right\rangle \Big|_{\sigma=1} - \left\langle \frac{d\rho(\sigma)}{d\sigma}, \nabla \Sigma_V \right\rangle \Big|_{\sigma=0} \geq K W_2^2(\rho(0), \rho(1)). \quad (35)$$

Let $\rho_t(s, x)dx : [0, 1] \rightarrow \mathcal{P}_2(\mathbb{R})$ be the unique geodesic between μ_t and μ_V . By Otto [32], we have the following derivative formula of the Wasserstein distance

$$\begin{aligned} \frac{d}{dt} W_2^2(\mu_t, \mu_V) &= -2 \int_{\mathbb{R}} \left\langle \frac{d\rho_t(s)}{ds}(x), \xi_t \right\rangle \Big|_{s=0} d\mu_t + 2 \int_{\mathbb{R}} \left\langle \frac{d\rho_t(s)}{ds}(x), \xi_t \right\rangle \Big|_{s=1} d\mu_V \\ &= 2 \left(- \left\langle \frac{d\rho_t(s)}{ds}(x), \nabla \Sigma_V \right\rangle \Big|_{s=1} + \left\langle \frac{d\rho_t(s)}{ds}(x), \nabla \Sigma_V \right\rangle \Big|_{s=0} \right) \\ &\leq -2K W_2^2(\mu_t, \mu_V), \end{aligned} \tag{36}$$

where in the last step we have used (35). The Gronwall inequality yields

$$W_2^2(\mu_t, \mu_V) \leq e^{-2Kt} W_2^2(\mu_0, \mu_V).$$

Recall that

$$\nabla \Sigma_V(\mu_V) = 0.$$

By the fact that μ_t is the gradient flow of Σ_V on $\mathcal{P}_2(\mathbb{R})$ and using the uniform K -convexity of Σ_V , we have

$$\begin{aligned} \frac{d}{dt} \|\nabla \Sigma_V(\mu_t)\|_{\mathcal{P}_2(\mathbb{R})}^2 &= 2 \left\langle \nabla \|\nabla \Sigma_V(\mu_t)\|_{\mathcal{P}_2(\mathbb{R})}^2, \frac{d\mu_t}{dt} \right\rangle \\ &= -2 \text{Hess} \Sigma_V(\mu_t) \left(\frac{d\mu_t}{dt}, \frac{d\mu_t}{dt} \right) \\ &\leq -2K \left\| \frac{d\mu_t}{dt} \right\|_{\mathcal{P}_2(\mathbb{R})}^2 \\ &= -2K \|\nabla \Sigma_V(\mu_t)\|_{\mathcal{P}_2(\mathbb{R})}^2. \end{aligned}$$

Since $\Sigma_V(\mu_V) = 0$, we derive that

$$\begin{aligned} \frac{d}{dt} \Sigma_V(\mu_t | \mu_V) &= \left\langle \nabla \Sigma_V(\mu_t), \frac{d\mu_t}{dt} \right\rangle \\ &= -\|\nabla \Sigma_V(\mu_t)\|_{\mathcal{P}_2(\mathbb{R})}^2 \\ &= \int_t^\infty \frac{d}{ds} \|\nabla \Sigma_V(\mu_s)\|_{\mathcal{P}_2(\mathbb{R})}^2 ds \\ &\leq -2K \int_t^\infty \|\nabla \Sigma_V(\mu_s)\|_{\mathcal{P}_2(\mathbb{R})}^2 ds \\ &= 2K \int_t^\infty \frac{d}{ds} \Sigma_V(\mu_s) ds \\ &= -2K \Sigma_V(\mu_t | \mu_V), \end{aligned}$$

which implies

$$\Sigma_V(\mu_t | \mu_V) \leq e^{-2Kt} \Sigma_V(\mu_0 | \mu_V).$$

The proof of Theorem 1.4 (ii) is completed. \square

To prove Theorem 1.4 (iii), we need the following free logarithmic Sobolev inequality and free Talagrand transportation cost inequality due to Ledoux and Popescu [25].

Theorem 4.4 (Ledoux-Popescu [25]) Suppose that V is a C^2 , convex and there exists a constant $r > 0$ such that

$$V''(x) \geq K > 0, \quad |x| \geq r.$$

Then there exists a constant $c = C(K, r) > 0$ such that the free Log-Sobolev inequality holds: for all probability measure μ with $I_V(\mu) < \infty$,

$$\Sigma_V(\mu|\mu_V) \leq \frac{2}{c} I_V(\mu).$$

Moreover, the free Talagrand transportation inequality holds: there exists a constant $C = C(K, r, V) > 0$ such that

$$CW_2^2(\mu, \mu_V) \leq \Sigma_V(\mu|\mu_V).$$

Proof of Theorem 1.4 (iii). By Biane and Speicher [4], we have the following entropy dissipation formula

$$\frac{\partial}{\partial t} \Sigma_V(\mu_t|\mu_V) = -\frac{1}{2} I_V(\mu_t).$$

By Theorem 4.4, there exists a constant $C_1 > 0$ such that the free LSI holds

$$\Sigma_V(\mu|\mu_V) \leq \frac{2}{C_1} I_V(\mu),$$

which yields

$$\frac{d}{dt} \Sigma_V(\mu_t|\mu_V) \leq -\frac{C_1}{4} \Sigma_V(\mu_t).$$

By the Gronwall inequality, we have

$$\Sigma_V(\mu_t|\mu_V) \leq e^{-C_1 t/4} \Sigma_V(\mu_0|\mu_V).$$

By Theorem 4.4 again, there exists a constant $C_2 > 0$ such that the free transportation cost inequality holds

$$W_2^2(\mu, \mu_V) \leq \frac{1}{C_2} \Sigma_V(\mu|\mu_V).$$

Therefore

$$W_2^2(\mu_t, \mu_V) \leq \frac{e^{-C_1 t/4}}{C_2} \Sigma_V(\mu_0|\mu_V).$$

This finishes the proof of Theorem 1.3 (iii). \square

5 Proof of Theorem 1.3 (ii) and (iii)

Proof of Theorem 1.3 (ii). In the proof of Theorem 1.4, we have proved the following inequalities

$$\Sigma_V(\mu_2(t)) - \Sigma_V(\mu_1(t)) \geq \left\langle \text{grad}_W \Sigma_V(\rho_t(s)), \frac{\partial}{\partial s} \rho_t(s) \right\rangle \Big|_{s=0} + \frac{K}{2} W_2^2(\mu_1(t), \mu_2(t)),$$

and

$$\Sigma_V(\mu_1(t)) - \Sigma_V(\mu_2(t)) \geq - \left\langle \text{grad}_W \Sigma_V(\rho_t(s)), \frac{\partial}{\partial s} \rho_t(s) \right\rangle \Big|_{s=1} + \frac{K}{2} W_2^2(\mu_1(t), \mu_2(t)).$$

Summing them together, we obtain

$$\left\langle \text{grad}_W \Sigma_V(\rho_t(s)), \frac{\partial}{\partial s} \rho_t(s) \right\rangle \Big|_{s=0} - \left\langle \text{grad}_W \Sigma_V(\rho_t(s)), \frac{\partial}{\partial s} \rho_t(s) \right\rangle \Big|_{s=1} \leq -K W_2^2(\mu_1(t), \mu_2(t)).$$

By Otto [32], we have the following derivative formula of the Wasserstein distance

$$\begin{aligned} \frac{d}{dt} W_2^2(\mu_1(t), \mu_2(t)) &= -2 \int_{\mathbb{R}} \left\langle \frac{d\rho_t(s)}{ds}(x), \xi_t \right\rangle \Big|_{s=0} d\mu_1(t) + 2 \int_{\mathbb{R}} \left\langle \frac{d\rho_t(s)}{ds}(x), \xi_t \right\rangle \Big|_{s=1} d\mu_2(t) \\ &= 2 \left(- \left\langle \frac{d\rho_t(s)}{ds}(x), \nabla \Sigma_V(\mu_1(t)) \right\rangle \Big|_{s=1} + \left\langle \frac{d\rho_t(s)}{ds}(x), \nabla \Sigma_V(\mu_2(t)) \right\rangle \Big|_{s=0} \right) \\ &\leq -2K W_2^2(\mu_1(t), \mu_2(t)), \end{aligned}$$

which implies

$$W_2(\mu_1(t), \mu_2(t)) \leq e^{-Kt} W_2(\mu_1(0), \mu_2(0)).$$

As a consequence, the McKean-Vlasov equation (5) has a unique weak solution. This finishes the proof of Theorem 1.3 (ii). \square

Proof of Theorem 1.3 (iii). By Theorem 1.3 (i), the family $\{L_N(t), t \in [0, T]\}$ is tight with respect to the weak convergence topology on $\mathcal{P}(\mathbb{R})$, and the limit of any convergent subsequence of $\{L_N(t), t \in [0, T]\}$ is a weak solution of (5). By the uniqueness of the weak solution to (5), we conclude that $L_N(t)$ weakly converges to $\mu(t)$. The proof of Theorem 1.3 (iii) is completed. \square

By the same argument as used in Otto [32] and Otto-Villani [33], we can prove the following HWI inequalities. To save the length of the paper, we omit the proof.

Theorem 5.1 (*HWI inequalities*) Suppose that there exists a constant $K \in \mathbb{R}$ such that

$$V''(x) \geq K, \quad \forall x \in \mathbb{R}.$$

Let $\mu_i \in \mathcal{P}_2(\mathbb{R})$, $i = 1, 2$. Then for all $t > 0$, we have

$$\Sigma_V(\mu_1) - \Sigma_V(\mu_2) \leq W_2(\mu_1, \mu_2) \|\text{grad}_W \Sigma_V(\mu_1)\|_{\mathcal{P}_2(\mathbb{R})} - \frac{K}{2} W_2^2(\mu_1, \mu_2). \quad (37)$$

In particular, for any solution to the McKean-Vlasov equation (5), we have

$$\Sigma_V(\mu(t)) \leq W_2(\mu(t), \mu_V) \|\text{grad}_W \Sigma_V(\mu(t))\|_{\mathcal{P}_2(\mathbb{R})} - \frac{K}{2} W_2^2(\mu(t), \mu(t)). \quad (38)$$

where

$$\|\text{grad}_W \Sigma_V(\rho)\|_{\mathcal{P}_2(\mathbb{R})}^2 = \int_{\mathbb{R}} \rho |V'(x) - 2H\rho(x)|^2 dx.$$

6 Double-well potentials and some conjectures

Theorem 1.3 and Theorem 1.4 apply to $V(x) = a|x|^p$ with $a > 0$ and $p \geq 2$. When $a = \frac{1}{2}$, $p = 2$ and $\beta = 1, 2, 4$, it corresponds to the GUE, GOE and GSE. Moreover, Theorem 1.3 and Theorem 1.4 also apply to the Kontsevich-Penner model on the Hermitian random matrices ensemble with external potential (cf. [10])

$$V(x) = \frac{ax^4}{12} - \frac{bx^2}{2} - c \log |x|.$$

Note that, for all $x \neq 0$,

$$\begin{aligned} V''(x) &= ax^2 + \frac{c}{x^2} - b \\ &\geq 2\sqrt{ac} - b > 0 \end{aligned}$$

provided that $a > 0$, $c > 0$ and $4ac > b^2$.

Let us consider the double well potential

$$V(x) = \frac{1}{4}x^4 + \frac{c}{2}x^2, \quad x \in \mathbb{R},$$

where $c \in \mathbb{R}$ is a constant. By [20, 6], it has been known that the density function of the equilibrium measure μ_V is explicitly given as follows:

(i) When $c < -2$,

$$\begin{aligned} \rho(x) &= \frac{1}{2\pi} |x| \sqrt{(x^2 - a^2)(b^2 - x^2)} & a < |x| < b, \\ &= 0 & \text{otherwise,} \end{aligned}$$

where $a^2 = -2 - c$ and $b^2 = 2 - c$.

(ii) When $c = -2$, $\rho(x) = \frac{1}{2\pi} x^2 \sqrt{4 - x^2}$ for $x \in [-2, 2]$ and $\rho(x) = 0$ otherwise.

(iii) When $c > -2$,

$$\begin{aligned} \rho(x) &= \frac{1}{\pi} (b_2 x^2 + b_0) \sqrt{a^2 - x^2} & |x| < a, \\ &= 0 & \text{otherwise,} \end{aligned}$$

where $a^2 = \frac{\sqrt{4c^2+48}-2c}{3}$, $b_0 = \frac{c+\sqrt{\frac{c^2}{4}+3}}{3}$, and $b_2 = \frac{1}{2}$.

When $c \in [0, \infty)$, V is C^2 convex and $V''(x) \geq 3$ for $|x| \geq 1$. In this case, Theorem 1.3 (iii) implies that $W_2(\mu_t, \mu_V) \rightarrow 0$ with exponential convergence rate. When $c < -2$, μ_V has two supports $[-b, -a]$ and $[a, b]$ which are disjoint. By Section 7.1 in Biane-Speicher [4], it is known that μ_t does not converge to μ_V . See also Biane [3] and Remark 1.7. This also indicates that one cannot simultaneously prove a free version of the Holley-Stroock logarithmic Sobolev inequality and a free version of the Talagrand T_2 -transportation cost inequality under bounded perturbations of the distribution of eigenvalues $p_N(dx) = Z_N^{-1} \prod_{i < j} |x_i - x_j|^2 \prod_{i=1}^N e^{-NV(x_i)} dx$. Otherwise, by analogue of the proof of

Theorem 1.4 (iii), we may prove that μ_t converges to μ_V with respect the W_2 -Wasserstein distance and hence in the weak convergence topology on $\mathcal{P}(\mathbb{R})$. See also [25, 26] for a discussion on non-convex potentials.

In the case $c \in [-2, 0)$, as the global minimizer μ_V of Σ_V has a unique support, and all stationary point of μ_V must satisfy the Euler-Lagrange equation $H\mu = \frac{1}{2}V'$, one can see that the Voiculescu free entropy Σ_V has a unique minimizer μ_V . As μ_t is the gradient flow of Σ_V on $\mathcal{P}_2(\mathbb{R})$, and since $\frac{d}{dt}\Sigma_V(\mu_t) = -2 \int_{\mathbb{R}} [V'(x) - 2(H\rho)(x)]^2 \rho(x) dx$, we see that $\Sigma_V(\mu_t)$ is strictly decreasing in time t unless μ_t achieves the (unique) minimizer μ_V . This yields that $\Sigma_V(\mu_t)$ converges to some value. The question whether $W_2(\mu_t, \mu_V) \rightarrow 0$ or even μ_t weakly converges to μ_V as $t \rightarrow \infty$ for the above double-well potential V remains open. We would like to raise the following conjectures.

Conjecture 6.1 *Consider the double-well potential $V(x) = \frac{1}{4}x^4 + \frac{c}{2}x^2$ with $c \in [-2, 0)$. Then μ_t converges to μ_V with respect the W_2 -Wasserstein distance and hence in the weak convergence topology on $\mathcal{P}(\mathbb{R})$.*

In general, we may raise the following conjectures.

Conjecture 6.2 *Suppose that the potential V is a C^2 potential function with $V''(x) \geq K_1$ for all $|x| \geq r$ and $V''(x) \geq -K_2$ for all $|x| \leq r$, where $K_1, K_2, r > 0$ are some constants. Suppose further that Σ_V has a unique minimizer which has a unique compact support. Then $\mu(t)$ converges to μ_V with respect the W_2 -Wasserstein distance and in the weak convergence topology on $\mathcal{P}(\mathbb{R})$.*

Conjecture 6.3 *Suppose that the potential V is a C^2 potential function with $V''(x) \geq -K$ for all $x \in \mathbb{R}$, where $K > 0$ is a constant. Suppose further that Σ_V has a unique minimizer which has a unique compact support. Then $\mu(t)$ converges to μ_V with respect the W_2 -Wasserstein distance and in the weak convergence topology on $\mathcal{P}(\mathbb{R})$.*

Finally, let us mention the following conjecture due to Biane and Speicher [4].

Conjecture 6.4 *Consider the double-well potential given by $V(x) = \frac{1}{2}x^2 + \frac{g}{4}x^4$, where g is a negative constant but very close to zero. Then μ_t weakly converges to μ_V .*

References

- [1] L. Ambrosio, N. Gigli, G. Savr , *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Birkh user-Verlag, Berlin, 2005.
- [2] G. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University Press., 2010.
- [3] P. Biane, Logarithmic Sobolev inequalities, matrix models and free entropy, Acta. Math. Sin. (Engl. Ser.) **19** (3) (2003), 497-506.
- [4] P. Biane, R. Speicher, Free diffusions, free energy and free Fisher information, Ann. Inst. H. Poincar  Probab. Stat. **37** (2001), 581-606.
- [5] D. Benedetto, E. Caglioti, J. A. Carrillo, M. Pulvirenti, A Non-Maxwell Steady distribution for one-dimensional granular media, J. Stat. Phys. 91 (1998), No5/6, 979-990.

- [6] P. Bleher, A. Its, Double Scaling Limit in the Random Matrix Model: The Riemann-Hilbert Approach, *Comm. Pure Appl. Math.* **56** (4) (2003), 433-516.
- [7] A. Boutet de Monvel, L. Pastur, M. Shcherbina, On the statistical mechanics approach in the random matrix theory: Integrated Density of States, *Journal of Statistical Physics*. Vol. 79. Nos. 3/4. 1995.
- [8] G. Blower, *Random Matrices: High Dimensional Phenomena*, London Math. Soc., Lect. Notes Ser. 367, Cambridge Univ. Press, 2009.
- [9] E. Brezin, A. Zee, Universality of the corrections between eigenvalues of large random matrices, *Nucl. Phys. B* 402, 613-627 (1993).
- [10] L. Chekhov, Yu. Makeenko, The multicritical Kontsevich-Penner model, *Morden Phys. Letters A*. Vol. **7** No. 14 (1992), 1223-1236.
- [11] J. Carrillo, R. McCann, C. Villani, Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates, *Rev. Mat. Iberoamericana* **19** (2003), 971-1018.
- [12] T. Chan, The Wigner semi-circle law and eigenvalues of matrix-valued diffusions. *Probab.Theory Relat.Fields* 93, 249-272 (1992).
- [13] P. Deift, *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach*, American Mathematical Society, 2000.
- [14] F. J. Dyson, Statistical theory of the energy levels of complex systems. I, II, and III. *J. Math. Phys.* 3. 140-156, 157-165, 166-175, 1962.
- [15] F. J. Dyson, A Brownian-motion model of the eigenvalues of a random matrix. *J. Mathematical Phys.* 3. 1191-1198, 1962.
- [16] S.N. Ethier, T.G. Kurtz, *Markov Processes: Characterization and convergence*, New York, Wiley 1986.
- [17] R. Fernandez, J. Fröhlich, A. Sokal, *Random Walks, Critical Phenomena, and Triviality in Quantum Field Theory*, Springer-Verlag, Heidelberg, 1992.
- [18] A. Guionnet, *Large random matrices: Lectures on macroscopic asymptotics*, Springer, 2008.
- [19] N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Publishing Company, Amsterdam, 1981.
- [20] K. Johansson, On fluctions of eigenvalues of random Hermitian matrices, *Duke Math. J.* **91** (1998), 151-204.
- [21] M. Kontsevich, Intersection theory on the moduli space of curves and the matrix Airy function, *Commun. Math. Phys.* 147 (1992), 1-23.
- [22] M. Katori, H. Tanemura, Non-equilibrium dynsamics of Dyson's model with an infinite number of particles, preprint, 2009.

- [23] C. Lancellotti, On the fluctuations about the Vlasov limit for N -particle systems with mean-field interactions, *J. Stat. Phys.* **136** (2009), 643-665.
- [24] B. Kloeckner, A geometric study of Wasserstein spaces: Euclidean spaces, *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze IX*, **2** (2010) 297-323, DOI10.2422/2036-2145.2010.2.03, and [www.arXiv:0804.3505](http://www.arXiv.org/abs/0804.3505).
- [25] M. Ledoux, I. Popescu, Mass transportation proofs of free functional inequalities, and free Poincaré inequalities, *J. Funct. Anal.* **257** (2009), 1175-1221.
- [26] M. Maïda, E. Maurel-Segala, Free transport-entropy inequalities for non-convex potentials and application to concentration for random matrices, hal-00687686 version 2, September 10, 2012.
- [27] F. Malrieu, Convergence to the equilibrium for granular media equations and their Euler schemes, *Ann. Appl. Probab.* **13** (2003), No. 2, 540-560.
- [28] M. Mehta, *Random Matrices*, 3rd Edition, Elsevier (Singapore) Pte Ltd, 2006.
- [29] U.F. Mayer, Gradient flows on non positively curved metric spaces and harmonic maps, *Commun. in Analysis and Geome.* **6**, (1998), No. 2, 199-253.
- [30] J. R. Norris, L.C.G. Rogers, D. Williams, Brownian motions of ellipsoids, *Trans. Amer. Math. Soc.* **294**, 2 (1986), 757-765.
- [31] H. Osada, Interacting Brownian motions in infinite dimensions with logarithmic interaction potentials, [arXiv:0902.3561v2](http://arXiv.org/abs/0902.3561)
- [32] F. Otto, The geometry of dissipative evolution equations: the porous medium equation, *Commun. Partial Differential Equations* **26** (2001), No.1/2, 101-174.
- [33] F. Otto, C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* **173** (2000) 361-400.
- [34] L. Pastur, M. Shcherbina, *Eigenvalue Distribution of Large Random Matrices*, *Amer. Math. Soc., Mathematical Surveys and Monographs*, Vol 171, 2010.
- [35] C. Prevot, M. Rockner, *A Concise Course on Stochastic Partial Differential Equations*, *Lecture Notes in Math.* 1905, Springer-Verlag, Berlin, 2007.
- [36] L.C.G. Rogers, Z. Shi, Interacting brownian particles and the Wigner law, *Probab. Theory Related Fields* **95**,4(1993), 555-570.
- [37] H. Spohn, *Large Scale Dynamics of Interacting Particles*, Springer-Verlag, Berlin Heidelberg 1991.
- [38] T. Tao, *Topics in Random Matrices Theory*, *Graduate Studies in Mathematics*, Vol. 132, *Amer. Math. Soc.* 2012.
- [39] C. Villani, *Topics in Mass Transportation*, *Grad. Stud. Math.*, *Amer. Math. Soc.*, Providence, RI, 2003.
- [40] C. Villani, *Optimal Transport, Old and New*, Springer-Verlage, Berlin, 2009.

- [41] D. Voiculescu, The analogues of entropy and of Fishers information measure in free probability theory, I, Commun. Math. Phys. 155 (1993), 71-92.
- [42] E. P. Wigner, On the distribution of the roots of certain symmetric matrices. Ann. of Math. (2) **67** (1958), 325-327.

Songzi Li, School of Mathematical Science, Fudan University, 220, Handan Road,
Shanghai, 200432, China

Xiang-Dong Li, Academy of Mathematics and Systems Science, Chinese Academy of
Sciences, 55, Zhongguancun East Road, Beijing, 100190, China
E-mail: xdli@amt.ac.cn

Yong-Xiao Xie, Academy of Mathematics and Systems Science, Chinese Academy of
Sciences, 55, Zhongguancun East Road, Beijing, 100190, China